Synopsis for Registration as Research

Student For

DoctoralDegree

On

Design of Novel Algorithms for Image Restoration using Evolutionary Techniques

Swati Ganeshlal Oza

Research Scholar

Supervisor

Dr. Archana Harsing Sable

SchoolofComputationalSciences,

SwamiRamanandTeerthMarathwadaUniversity,NandedN

anded-431606,MaharashtraState(MS),India.

Abstractofproposedworkplan

Conventional methods for solving image restoration problems are typically built on an image degradation model and on some priors of the latent image. The model of the degraded image and the prior knowledge of the latent image are necessary because the restoration is an ill posted inverse problem. However, for some applications, such as those addressed in this thesis, the image degradation process is too complex to model precisely; in addition, mathematical priors, such as low rank and sparsity of the image signal, are often too idealistic for real world images. These difficulties limit the performance of existing image restoration algorithms, but they can be, to certain extent, overcome by the techniques of machine learning, particularly deep convolutional neural networks. Machine learning allows large sample statistics far beyond what is available in a single input image to be exploited. More importantly, the big data can be used to train deep neural networks to learn the complex non-linear mapping between the degraded and original images. This circumvents the difficulty of building an explicit realistic mathematical model when the degradation causes are complex and compounded.

In this synopsis, we design and implement deep convolution neural networks (DCNN) for two challenging image restoration problems: reflection removal and joint demosaicking-deblurring. The first problem is one of blind source separation; its DCNN solution requires a large set of paired clean and mixed images for training. As these paired training images are very difficult, if not impossible, to acquire in the real world, we develop a novel technique to synthesize the required training images that satisfactorily approximate the real ones. For the joint demosaicking-deblurring problem, we propose a new multiscale DCNN architecture consisting of a cascade of sub networks so that the underlying blind deconvolution task can be broken into smaller sub problems and solved more effectively and robustly. In both cases extensive experiments are carried out. Experimental results demonstrate clear advantages of the proposed DCNN methods over existing ones.

1. Introduction

Image restoration is a process to recover high quality clean images from their degraded counterparts. The degradation can have a variety of causes, such as sensor noises, down sampling, motion blur, reflection and some combinations of the above. Restoring an input image to its best form is a necessary and important step in many applications of image processing and low-level computer vision. There is a large body of literature of image restoration, covering techniques of denoising, super resolution, deblurring, demosaicking, reflection removal, etc.

Main stream image restoration methods are based on image degradation models and on some priors of the latent image. However, due to the complexity of real-world image degradation, it is often extremely difficult to model the process precisely. In addition, image restoration is in general a severely underdetermined inverse problem, as a large number of different latent images can degrade into the same image. Solving the inverse problem requires additional priors, such as low rank and sparsity of the image signal, to restrict the solution space. However, the priors are necessarily mathematical constructs that are often too idealistic for real-world images. These difficulties greatly limit the performance of the conventional image restoration algorithms.

The newly emerged data-driven machine learning method, particularly deep convolutional neural networks, open up new possibilities for more powerful and versatile solutions of image restoration. In machine learning one can use very large sample data set to make statistical inferences that is much more accurate and robust than relying only on the observations of a single input image. Also, big data facilitates the training of deep neural networks to learn highly complex non-linear mapping between the degraded and original images. This makes it possible to solve the inverse problem of image restoration without building an explicit realistic mathematical model. Indeed, past few years have witnessed significant improvements of image restoration results made by deep learning over traditional image processing methods.

Artificial neural network (ANN) is a widely used method of machine learning that mimics how the biological neural network processes information in human brain. It can solve difficult problems in diverged areas of applications, such as computer vision, natural language processing, text processing, etc. ANNs consist of layers of neurons, each neuron has a weight associated with it. A neuron is a basic computation unit; it multiplies the weight with its input, and computes the output with a non-linear activation function. The networks are optimized by back propagation algorithms to minimize a loss function, which is, in the case of image restoration, the difference between the restored and ground truth images.

During training, the network performs forward and backward propagation alternatively to evaluate the loss function and update the parameters. The loss function is the key of training and is carefully designed so that the training process can be seen as an optimization problem. Besides the loss function, training data also affects the performance of the network. Insufficient training data and statistically mismatched training data will cause over-fitting and non-convergence of the network. By increasing the number of hidden layers, the neural network can fit more complex functions and it also requires more training data.

With the invention of convolutional neural network (CNN), many attempts have been made to address different computer vision problems. CNNs are multi-layer structure neural network and they are aimed to extract complicated information from the input images. They are trained to learn high-level discriminative features in an incremental way for classifying structured data such as images and videos. CNN architecture consists of convolutional, nonlinear, pooling and fully connected layers. Each convolutional layer applies convolution on its input tensor with its filters, each filter generates feature maps with an non-linear activation function. Higher level convolutional layers extract more abstract features that contains discriminative information for making the final decision.

Image restoration is process of recovering the original image by removing noise and blur from image. Image blur is difficult to avoid in many situations like photography, to remove motion blur caused by camera shake, radar imaging to remove the effect of image system response, etc. Image noise is unwanted signal which comes in image from sensor such as thermal or electrical signal and Environmental condition such as rain, snow etc. Researchers have proposed many methods in this regard and in this paper we will examine and discuss different noise and blur models and restoration methods.

The first successful deep neural network is the well knownAlexNet published in 2012 [9]. It made a breakthrough in visual object recognition and achieved unprecedented accuracy. The power of AlexNet largely comes from its adoption of convolutional kernels in first few layers of the network, thus this type of network is called convolutional neural network (CNN).

It consists of five stacked convolutional layers followed by two fully connected layers. Convolutional layers extract features, while fully connected layers estimate the probabilities of the output classes based on the features. Shortly after the AlexNet, deep learning has become an extremely active research field, drawing attentions from almost all academic disciplines. Thanks to the rapid advances of high performance computation hardware, especially GPUs, larger and larger network models can be built. Many researchers have been trying to make CNNs more powerful by increasing the network depth and width. Based on AlexNet, VGG net is proposed [10] that increases the number of convolutional layers while using smaller 3×3 kernels. The increase in network depth boosts the performance significantly, reducing the classification error by a large margin.

In addition to the successes of deep learning techniques in image recognition, image restoration can also be performed by CNN-based deep learning with the quality of restored images far exceeding that of traditional methods. A straightforward way of deep learning for image restoration is to remove the fully connected layers of a CNN.

classification network. Early methods simply stack convolution layers to perform mapping between degraded and latent images with respect to a specific restoration problem [11]. A more effective approach is to use deep convolutional encoder-decoder networks, first proposed by Mao et al. [12]. In many applications, the degraded image and its clean version are very close, it will be more efficient to only learn the difference between them. Hence, stacking residual blocks becomes a popular technique and it generates superior results [13, 5].

2. Review of Literature and Development in the subject

Many existing reflection removal methods rely on two or more input images of the same scene with different reflections to estimate the transmission layer. To obtain such images of varied reflections, several photography techniques can be used. Some reflection removal methods employ polarizing filter to manipulate the level of the reflections [22, 23, 24]. The physically-based method proposed by Schechner et al. shows the advantages of using orthogonal polarized input images [25]. Kong et al. further improve this idea by exploiting the spatial properties of polarization [15]. 11 M.A.Sc. Thesis - Zhixiang Chi McMaster - Electrical Engineering Similar to polarizing filtering, flash lighting [26, 27] and defocus blurring [28, 29] can help generate images with varied reflections without moving the camera as well. Keeping the camera relatively stationary is crucial to the efficacy of these device-based reflection removal techniques, as if the objects behind the glass are also not in motion, the only changing components among the input images are the reflections while the transmission layer is invariant.

There are also many multi-image techniques that exploit the motion of the camera as a cue for reflection removal. These techniques first align the objects in a series of images taken from slightly different viewing positions and then separate the invariant layer as the reflection-free image [30, 16]. To align images interfered by reflections, Tsinet at. [31] andSinha et al. [32] use efficient stereo matching algorithms. Guo et al. exploit the sparsity and independence of the transmission and reflection layers to improve the robustness of image alignment [33]. Off-the-shelf optical flow algorithms are also employed for aligning images [17, 34]. With motion smoothness constraints, optical flow techniques can be more accurate and robust for the layer separation task [35, 36]. For the cases where the camera is stationary while the objects are moving, the reflection layer is relatively static and must be handled differently [37, 38].

If only one image of the scene is given, which is the case tackled by this thesis, the task of reflection removal becomes much more challenging. Only a few single image reflection removal techniques have been reported in the literature. Many of these techniques still rely on some extra information provided by light field camera [39, 40] or the user [41, 42]. One of the first attempts to solve the problem without any user assistance is [18], which minimizes the total amount of edges and corners in the two decomposed layers of the input image. Akashi et al. [43] employ sparse non-negative matrix factorization to separate the reflection layer without a explicit smoothness prior. The work of Li and Brown [21] assumes that the reflection layer is smoother 12 M.A.Sc. Thesis - Zhixiang Chi McMaster - Electrical Engineering than the transmission layer due to defocus blur and hence has a short tail gradient distribution. With a similar smoothness assumption for the reflection layer, Fan et al. [19] use two cascaded CNN networks to reconstruct a reflection reduced image from the edges of the input image. Arvanitopoulos et al. [44] formulate the reflection suppression problem as an optimization problem with a Laplacian data fidelity term and a total variation term. Wan et al. [45] combine the sparsity prior and nonlocal prior of image patches in both the transmission and reflection layers together. They further increase the effectiveness of the nonlocal prior using image patches retrieved from an external dataset.

The work of Shih et al. [20] takes advantage of ghosting, the phenomenon of multiple reflections caused by thicker glass, and decomposes the input image based on Gaussian mixture model (GMM). To deal with reflections from eyeglasses in frontal face image, Sandhan and Choi [46] exploit the bilateral symmetry of human face and use a mirrored input image as another input of varied reflections.For more detailed review on the existing techniques for reflection removal, we refer readers to two excellent surveys [47] and [48].

3. Objectives of Research

In the literature survey, various deep learning techniques used in image restorations are discussed. Convolutional neural networktechniques, Classic Neural Networkstechniques, Recurrent Neural Networkstechniques are discussed. The available restoration methods havesome limitation.

Listed below are the limitations of the Current Image Restoration and Classification methods:

1. To introduce novel hybrid deep learning technique for Image Restoration using convolutional neural network.
2. To study and analyze the image pre-processing algorithms to enhance the quality of input image.
3. To Introduce Deep Learning Techniques for Image Restoration.
4. To implement deep convolutional neural networks.
5. To implement reflection removal method for image Restoration.

4. Methodology to be Adopted

Workplan:

The proposed network consists of 12 convolutional layers and 12 deconvolutional with one rectified linear unit (ReLU) following each of the layers. The left graph of Figure ??shows how ReLU works.The convolutional layers are designed to extract and condense features from the input, while deconvolutional layers rebuild the details of reflection-free image from feature abstractions. Overall, the architecture of proposed encoder-decoder network can be divided into three stages:

- **Feature extraction**: The 6 convolutional layers in this stage extract features for both transmission and reflection layers.

- **Reflection recovery and removal** :In this stage, the first 6 convolutional layers and following 6 deconvolutional layers are set to learn and recover reflection. Additionally, to preserve the details of the reflection layer better, two skip connections are added in the second stage to inherit the features learned from previous convolutional layers [12]. by minimizing the error of the output against the ground-truth reflection, this stage preserves only the features of the reflections and removes the transmission layer gradually from the input. At the end of this stage, the extracted reflection features are removed before deconv 7, the seventh deconvolutional layer, by using an element-wise subtraction [51] followed by a ReLU activation max(0, conv 6−deconv 6).

- **Transmission layer restoration.** The reflections might not be removed completely after previous stage by simply subtracting the estimated reflection.Thus, this stage tries to restore a visually pleasing transmission image from the reflection subtracted image. To achieve this goal, 6 deconvolutional layers are used to recover transmission layer from the features of the targeted scene. For image classification tasks, pooling layers are necessary as it extracts main abstract features that are crucial for final decision [2]. However, as the redundant information increases the difficulty for deconvolutional layers to recover the image.

- Convolutional neural network is a class of deep learning methods which has become dominant in various computer vision tasks and is attracting interest across a variety of domains, including radiology.

- Convolutional neural network is composed of multiple building blocks, such as convolution layers, pooling layers, and fully connected layers, and is designed to automatically and adaptively learn spatial hierarchies of features through a backpropagation algorithm.

- Familiarity with the concepts and advantages, as well as limitations, of convolutional neural network is essential to leverage its potential to improve radiologist performance and, eventually, patient care.

5. Importance of Society application

Wetackle two image restoration problems: single image reflection removal and joint demosaicking and deblurring. For the former problem, we propose a new reflection formation model for generating synthetic photo-realistic images as training data, and we design a three stage convolutional encoder-decoder network to remove reflections.

The feature maps in each convolutional layer for the trained encoder-decoder network show that each stage is working as intended. For the second problem, we proposed recursive multi scale network to solve demosaicking and deblurring jointly in a coarse-to-fine manner. The network explores the information on four independent Bayer structures. Combining with a sub network that performs deblurring process recursively, the network is able to produce high fidelity images from blurry Bayer images. Extensive experimental results show that the proposed methods outperform other state of arts for both applications.

As a result, conventional approaches often perform poorly in real-world scenarios with complex or compound degradations or both. In comparison, the newly emerged data driven approach, DCNN, is potentially more robust for real-world image restoration problems. Without employing any explicit degradation models, DCNN based techniques learn the mapping from a degraded image to its corresponding clean image through a large number of such sample pairs.

6. Proposed work Plan

Theapproximatedurationproposedtocompleteresearchprojectin36months.Presentationwill be done after every 6 months

- Literaturereview-after6monthsandImplementationofbasicpreprocessingalgorithmsto improve the imagequality.
- ImplementationofavailableRestorationalgorithmsandproposedrestorationalgorithms.
- Implementationof deep learning techniques for Image Restoration.
- Comparingtheresultsobtainedfromproposedandexisting restorationalgorithm.

Finalsubmissionofthesiswillbeafterpermissionofguideandfulfillingminimumrequirements.

References

[1] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition. (2015) 1–9

[2] He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). (2016) 770–778

[3] Losson, O., Macaire, L., Yang, Y.: Comparison of color demosaicing methods. In: Advances in Imaging and Electron Physics. Volume 162. Elsevier (2010) 173–265

[4] Michaeli, T., Irani, M.: Blind deblurring using internal patch recurrence. In: European Conference on Computer Vision, Springer (2014) 783–798

[5] Nah, S., Kim, T.H., Lee, K.M.: Deep multi-scale convolutional neural network for dynamic scene deblurring. In: IEEE CVPR. Volume 2017. (2017)

[6] Lai, W.S., Huang, J.B., Hu, Z., Ahuja, N., Yang, M.H.: A comparative study for single image blind deblurring. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2016) 1701–1709

[7] Gharbi, M., Chaurasia, G., Paris, S., Durand, F.: Deep joint demosaicking and denoising. ACM Transactions on Graphics (TOG) 35(6) (2016) 191

[8] Malvar, H.S., He, L.w., Cutler, R.: High-quality linear interpolation for demosaicing of bayer-patterned color images. In: Acoustics, Speech, and Signal Processing, 2004. Proceedings.(ICASSP'04). IEEE International Conference on. Volume 3., IEEE (2004) iii–485

[9] Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems. (2012) 1097–1105

[10] Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)

[11] Dong, C., Loy, C.C., He, K., Tang, X.: Learning a deep convolutional network for image super-resolution. In: European conference on computer vision, Springer (2014) 184–199

[12] Mao, X.J., Shen, C., Yang, Y.B.: Image restoration using convolutional autoencoders with symmetric skip connections. arXiv preprint arXiv:1606.08921 (2016)

[13] Ledig, C., Theis, L., Husz´ar, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., et al.: Photo-realistic single image super-resolution using a generative adversarial network. arXiv preprint arXiv:1609.04802 (2016)

[14] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial nets. In: Advances in neural information processing systems. (2014) 2672–2680

[15] Kong, N., Tai, Y.W., Shin, J.S.: A physically-based approach to reflection separation: from physical modeling to constrained optimization. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI) 36(2) (2014) 209–221

[16] Gai, K., Shi, Z., Zhang, C.: Blind separation of superimposed moving images using image statistics. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI) 34(1) (2012) 19–32

[17] Li, Y., Brown, M.S.: Exploiting reflection change for automatic reflection removal. In: IEEE International Conference on Computer Vision (ICCV). (2013) 2432–2439

[18] Levin, A., Zomet, A., Weiss, Y.: Separating reflections from a single image using local features. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Volume 1. (2004) 306–313

[19] Fan, Q., Yang, J., Hua, G., Chen, B., Wipf, D.: A generic deep architecture for single image reflection removal and image smoothing. In: IEEE International Conference on Computer Vision.

[20] Shih, Y., Krishnan, D., Durand, F., Freeman, W.T.: Reflection removal using ghosting cues. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). (2015) 3193–3201
[21] Li, Y., Brown, M.S.: Single image layer separation using relative smoothness. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). (2014) 2752–2759
[22] Ohnishi, N., Kumaki, K., Yamamura, T., Tanaka, T.: Separating real and virtual objects from their overlapping images. In: European Conference on Computer Vision (ECCV). (1996) 636–646
[23] Farid, H., Adelson, E.H.: Separating reflections and lighting using independent components analysis. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Volume 1. (1999) 262–267
[24] Sarel, B., Irani, M.: Separating transparent layers through layer information exchange. In: European Conference on Computer Vision (ECCV). (2004) 328– 341
[25] Schechner, Y.Y., Shamir, J., Kiryati, N.: Polarization and statistical analysis of scenes containing a semireflector. Journal of the Optical Society of America A (JOSA A) 17(2) (2000) 276–284
[26] Feris, R., Raskar, R., Tan, K.H., Turk, M.: Specular reflection reduction with multi-flash imaging. In: Brazilian Symposium on Computer Graphics and Image Processing. (2004) 316–321
[27] Agrawal, A., Raskar, R., Nayar, S.K., Li, Y.: Removing photography artifacts using gradient projection and flash-exposure sampling. ACM Transactions on Graphics (TOG) 24(3) (2005) 828–835
[28] Schechner, Y.Y., Kiryati, N., Shamir, J.: Blind recovery of transparent and semireflected scenes. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Volume 1. (2000) 38–43
[29] Schechner, Y.Y., Kiryati, N., Basri, R.: Separation of transparent layers using focus. International Journal of Computer Vision (IJCV) 39(1) (2000) 25–39
[30] Szeliski, R., Avidan, S., Anandan, P.: Layer extraction from multiple images containing reflections and transparency. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Volume 1. (2000) 246–253
[31] Tsin, Y., Kang, S.B., Szeliski, R.: Stereo matching with linear superposition of layers. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI) 28(2) (2006) 290–301
[32] Sinha, S.N., Kopf, J., Goesele, M., Scharstein, D., Szeliski, R.: Image-based rendering for scenes with reflections. ACM Transactions on Graphics (TOG) 31(4) (2012) 100:1–100:10
[33] Guo, X., Cao, X., Ma, Y.: Robust separation of reflection from multiple images. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). (2014) 2187–2194
[34] Han, B.J., Sim, J.Y.: Reflection removal using low-rank matrix completion. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). (2017) 5438–5446
[35] Xue, T., Rubinstein, M., Liu, C., Freeman, W.T.: A computational approach for obstruction-free photography. ACM Transactions on Graphics (TOG) 34(4) (2015) 79:1–79:11
[36] Yang, J., Li, H., Dai, Y., Tan, R.T.: Robust optical flow estimation of doublelayer images under transparency or reflection. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). (2016) 1410–1419
[37] Sarel, B., Irani, M.: Separating transparent layers of repetitive dynamic behaviors. In: IEEE International Conference on Computer Vision (ICCV). Volume 1. (2005) 26–32
[38] Simon, C., Kyu Park, I.: Reflection removal for in-vehicle black box videos. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). (2015) 4231–4239
[39] Chandramouli, P., Noroozi, M., Favaro, P.: Convnet-based depth estimation, reflection separation and deblurring of plenoptic images. In: Asian Conference on Computer Vision (ACCV). (2016) 129–144
[40] Ni, Y., Chen, J., Chau, L.P.: Reflection removal based on single light field capture. In: IEEE International Symposium on Circuits and Systems (ISCAS). (2017) 1–4

[41] Levin, A., Weiss, Y.: User assisted separation of reflections from a single image using a sparsity prior. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI) 29(9) (2007) 1647–1654
[42] Yeung, S.K., Wu, T.P., Tang, C.K.: Extracting smooth and transparent layers from a single image. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). (2008)
[43] Akashi, Y., Okatani, T.: Separation of reflection components by sparse nonnegative matrix factorization. In: Asian Conference on Computer Vision (ACCV). (2014) 611–625
[44] ArvanitopoulosDarginis, N., Achanta, R., S¨usstrunk, S.: Single image reflection suppression. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). (2017)
[45] Wan, R., Shi, B., Tan, A.H., Kot, A.C.: Sparsity based reflection removal using external patch search. In: IEEE International Conference on Multimedia and Expo (ICME). (2017) 1500–1505
[46] Sandhan, T., Choi, J.Y.: Anti-glare: Tightly constrained optimization for eyeglass reflection removal. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). (2017) 1241–1250
[47] Artusi, A., Banterle, F., Chetverikov, D.: A survey of specularity removal methods. Computer Graphics Forum 30(8) (2011) 2208–2230
[48] Wan, R., Shi, B., Duan, L.Y., Tan, A.H., Kot, A.C.: Benchmarking single-image reflection removal algorithms. In: IEEE International Conference on Computer Vision (ICCV). (2017) 3922–3930
[49] Diamant, Y., Schechner, Y.Y.: Overcoming visual reverberations. In: Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, IEEE (2008) 1–8
[50] Kim, J., Kwon Lee, J., Mu Lee, K.: Accurate image super-resolution using very deep convolutional networks. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). (2016) 1646–1654

Swati GaneshlalOza
Research Scholar

Dr.ArchanaHarsingSable
Research Supervisor